

ЛЬВІВСЬКИЙ НАЦІОНАЛЬНИЙ УНІВЕРСИТЕТ ІМЕНІ ІВАНА ФРАНКА

На правах рукопису

Волощук Орест Романович

УДК 004.853+004.855.5

МЕТОДИ КЛАСТЕРИЗАЦІЇ НА ВЕЛИКИХ МАСИВАХ ДАНИХ

01.05.03 — математичне та програмне забезпечення обчислювальних
машин і систем

Магістерська робота

Науковий керівник

Годич Олесь Васильович,

кандидат фізико-математичних наук, доцент

Львів — 2011

ЗМІСТ

Вступ

3

ВСТУП

Актуальність теми. Сьогодні перед дослідниками та науковцями все частіше постають задачі аналізу даних та здійснення прогнозів чи прийняття рішень на підставі результатів аналізу. Із розвитком обчислювальної техніки збільшуються також об'єми баз даних, що піддаються аналізу. Виникає необхідність створення оптимальних алгоритмів обробки великих масивів даних. Одним із важливих та найбільш ресурсоемних етапів такої обробки є кластеризація.

Мета і завдання дослідження. Метою дослідження є розвиток методики кластеризації великих масивів даних. Для досягнення цієї мети були сформульовані та вирішені такі основні завдання:

- провести детальний аналіз та дослідити ефективність різноманітних алгоритмів кластеризації;
- здійснити контроль якості результатів роботи алгоритмів;
- розробити методи та виявити можливі оптимізації, що дозволять прискорити виконання задачі кластеризації

Об'єкт дослідження. Об'єктом дослідження є методи ієрархічні та плоскі алгоритми кластеризації даних.